

Second-Order Resolvability, Intrinsic Randomness, and Fixed-Length Source Coding for Mixed Sources: Information Spectrum Approach

Ryo Nomura, *Member, IEEE*, and Te Sun Han, *Life Fellow, IEEE*

Abstract

The second-order achievable asymptotics in typical random number generation problems such as resolvability, intrinsic randomness, fixed-length source coding are considered. In these problems, several researchers have derived the first-order and the second-order achievability rates for general sources using the information spectrum methods. Although these formulas are general, their computation are quite hard. Hence, an attempt to address explicit computation problems of achievable rates is meaningful. In particular, for i.i.d. sources, the second-order achievable rates have earlier been determined simply by using the asymptotic normality. In this paper, we consider mixed sources of two i.i.d. sources. The mixed source is a typical case of nonergodic sources and whose self-information does not have the asymptotic normality. Nonetheless, we can explicitly compute the second-order achievable rates for these sources on the basis of two-peak asymptotic normality. In addition, extensions of our results to more general mixed sources, such as a mixture of countably infinite i.i.d. sources or Markovian sources, and a continuous mixture of i.i.d. sources, are considered.

Index Terms

R. Nomura is with the School of Network and Information, Senshu University, Kanagawa, Japan, e-mail: nomu@isc.senshu-u.ac.jp

T. S. Han is with the National Institute of Information and Communications Technology (NICT), Tokyo, Japan, e-mail: han@is.uec.ac.jp

The first author with this work was supported in part by JSPS Grant-in-Aid for Young Scientists (B) No. 23760346.

I. INTRODUCTION

The problem of random number generation is one of the main topics in information theory [1]–[4]. There are several problem settings in random number generation. In particular, the resolvability problem and the intrinsic randomness problem are representative of them. The *resolvability* problem is formulated as follows [5]. We first use the term of “*general source*” to denote a sequence $\mathbf{X} = \{X^n\}_{n=1}^\infty$ of random variables X^n indexed by n (taking values in *countably infinite* sets), typically, n -dimensional random variables. Given an arbitrary general source $\mathbf{X} = \{X^n\}_{n=1}^\infty$ (called the *target* random number), we generate or approximate it by using a discrete *uniform* random number whose size is requested to be as small as possible. One of the main objectives in this problem is to construct an efficient algorithm that transforms the discrete uniform random number to the specified target source. Han and Verdú [1], and Steinberg and Verdú [6] have determined the infima of achievable uniform random number rates by using the information spectrum methods. On the other hand, the *intrinsic randomness* problem is formulated as follows in Vembu and Verdú [2]: Given an arbitrary general source $\mathbf{X} = \{X^n\}_{n=1}^\infty$ (called the *coin* source), we try to generate or approximate, by using $\mathbf{X} = \{X^n\}_{n=1}^\infty$, a uniform random number at as large rates as possible. Vembu and Verdú [2] and Han [5] have determined the suprema of achievable uniform random number generation rates, again by invoking the information spectrum methods. Since the class of *general* sources is quite large and it includes all nonstationary and/or nonergodic sources, their results are very basic and quite fundamental.

Furthermore, it turned out that these random number generation problems have close bearing with the fixed-length source coding problem (cf. Han and Verdú [1]). All the formulas established here may be said to be ones of the *first-order*.

On the other hand, the finer evaluation of achievable rates, called the *second-order* achievable rates, have been investigated in several contexts. In variable-length source coding problem, Kontoyiannis [7] has established the second-order source coding theorem. In channel coding problem, Strassen (see, Csiszár and Körner [8]), Hayashi [9], and Polyanskiy, Poor and Verdú [10] have determined the second-order capacity rates. Hayashi [11] has also shown the second-order achievability theorems for the intrinsic randomness problem as well as for the fixed-length source coding problem with *general* sources. In addition, for i.i.d. sources he has demonstrated the calculation of

the second-order optimal achievable rates by using the asymptotic normality in both problems.

In the present paper, we address the computation problem concerning the second-order formulas for resolvability, intrinsic randomness, and fixed-length source coding problems for *mixed* (non-i.i.d.) sources, where the resolvability problem was first and partly studied by Nomura and Matsushima [12]. In the resolvability problem or the intrinsic randomness problem the degree of approximation is measured in terms of variational distance.

As we have mentioned in the above, the analysis based upon the asymptotic normality is effective in deriving the second-order achievable rates. However, it had earlier been applied only to the class which has a simple probabilistic structure, such as i.i.d. sources, Markovian sources or stationary discrete memoryless channels but *not* to that of *mixed* sources.

In the present paper, we first establish the resolvability formula for general sources, and then specifically compute the second-order optimal achievable rates for *mixed* sources, which is a wider but still tractable class of sources than the previous ones in random number generation. Recall that mixed sources are typical cases of *nonergodic* sources. Nonetheless, we show that we can use still the *two-peak* asymptotic normality to compute the second-order achievable rates for mixed sources.

Related works include, e.g., Polyanskiy, Poor and Verdú [13] that has developed the second-order capacity of the Gilbert-Elliott channel (GEC). The GEC is a simple model of channels in which the crossover probability of a *binary* symmetric channel obeys a *binary* symmetric Markov chain transition. In particular, they have analyzed the *nonergodic* case in the GEC using the two-peak asymptotic normality from the viewpoint of a *mixture* of two memoryless channels. On the other hand, in this paper we consider a mixed source consisting of two i.i.d sources with *countably infinite* alphabet. Our analysis is on the basis of the *information spectrum* methods, and hence valid with countably infinite alphabet. In addition, it is shown that our results can be easily extended to more general cases. Actually, we generalize our results to mixed sources consisting of *countably infinite* i.i.d. sources with *countably infinite* alphabet, and furthermore to mixed sources consisting of *countably infinite* Markovian sources but with *finite* alphabet. The resolvability formula for mixed sources with *general mixture* not necessarily countably infinite mixture is also established.

It should be emphasized here that we have recourse to a bulk of *information spectrum calculations* throughout in the paper. Although they apparently look tedious and even cumbersome, each step in the process of computations is actually simple and very basic, which features the information spectrum unifying approach.

This paper is organized as follows. In Section II, we review the previous results on the first-order asymptotics for *general* sources. In section III, we review and derive the second-order asymptotic formulas for the *general* sources, analogously to Section II. In Section IV, we define the mixed source and state the lemmas which play the key role in the subsequent analysis. In Sections V, VI and VII, with *mixed* sources we establish the second-order achievability by invoking the *two-peak* asymptotic normality, for the resolvability problem, the intrinsic randomness problem and the fixed-length source coding problem, respectively. In Section VIII, we point out that all the results established in Sections V–VII are still valid if we consider more general mixed sources. Finally, we conclude our results in Section IX.

II. FIRST-ORDER ASYMPTOTICS

In this section we review the previous results on the first-order asymptotics of random number generation and fixed-length source coding.

To this end, we first give the necessary notations and definitions. In the sequel, let $\mathbf{Y} = \{Y^n\}_{n=1}^\infty$ be a general source with values in countable sets \mathcal{Y}^n . Let \mathcal{Z} be a countable set and let Z, \bar{Z} be random variables with values in \mathcal{Z} . Denote by $d(Z, \bar{Z})$ the variational distance

$$d(Z, \bar{Z}) \equiv \sum_{z \in \mathcal{Z}} |P_Z(z) - P_{\bar{Z}}(z)|,$$

where $P_X(\cdot)$ denotes the probability distribution of random variable X . Moreover, set $\mathcal{U}_M \equiv \{1, 2, \dots, M\}$ and let U_M denote the random variable uniformly distributed on \mathcal{U}_M .

A. First-Order Resolvability

Definition 2.1: Rate R is said to be δ -achievable if there exists a mapping $\phi_n : \mathcal{U}_{M_n} \rightarrow \mathcal{Y}^n$ such that

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log M_n \leq R \text{ and } \limsup_{n \rightarrow \infty} d(Y^n, \phi_n(U_{M_n})) \leq \delta.$$

Definition 2.2 (δ -resolvability):

$$S_r(\delta|\mathbf{Y}) = \inf \{R | R \text{ is } \delta\text{-achievable}\}.$$

Then, we have

Theorem 2.1 (Steinberg and Verdú [6]):

$$S_r(\delta|\mathbf{Y}) = \inf \left\{ R \left| F(R) \leq \frac{\delta}{2} \right. \right\} \quad (0 \leq \forall \delta < 2),$$

where

$$F(R) = \limsup_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n} \log \frac{1}{P_{Y^n}(Y^n)} \geq R \right\}. \quad (2.1)$$

The following Fig. 2.1 illustrates Theorem 2.1.

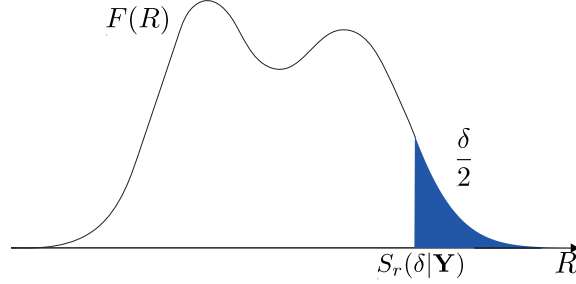


Fig. 2.1. First-Order Resolvability Rate

B. First-Order Intrinsic Randomness

Definition 2.3: Rate R is said to be δ -achievable if there exists a mapping $\phi_n : \mathcal{Y}^n \rightarrow \mathcal{U}_{M_n}$ such that

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log M_n \geq R \text{ and } \limsup_{n \rightarrow \infty} d(U_{M_n}, \phi_n(Y^n)) \leq \delta.$$

Definition 2.4 (δ -intrinsic randomness):

$$S_\iota(\delta|\mathbf{Y}) = \sup \{R \mid R \text{ is } \delta\text{-achievable}\}.$$

Then, we have

Theorem 2.2 (Han [5]):

$$S_\iota(\delta|\mathbf{Y}) = \sup \left\{ R \mid G(R) \leq \frac{\delta}{2} \right\} \quad (0 \leq \forall \delta < 2),$$

where

$$G(R) = \limsup_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n} \log \frac{1}{P_{Y^n}(Y^n)} \leq R \right\}.$$

The following Fig. 2.2 illustrates Theorem 2.2.

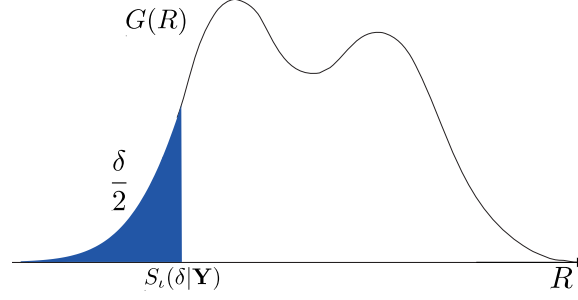


Fig. 2.2. First-Order Intrinsic Randomness Rate

C. First-Order Fixed-length Source Coding

Let $\varphi_n : \mathcal{Y}^n \rightarrow \mathcal{U}_{M_n}$, $\psi_n : \mathcal{U}_{M_n} \rightarrow \mathcal{Y}^n$ be an encoder and a decoder, respectively, for source $\mathbf{Y} = \{Y^n\}_{n=1}^\infty$. The decoding error probability ε_n is given by $\varepsilon_n \equiv \Pr \{Y^n \neq \psi_n(\varphi_n(Y^n))\}$. Such a code is denoted by (n, M_n, ε_n) .

Definition 2.5: Rate R is said to be ε -achievable if there exists a code (n, M_n, ε_n) such that

$$\limsup_{n \rightarrow \infty} \varepsilon_n \leq \varepsilon \text{ and } \limsup_{n \rightarrow \infty} \frac{1}{n} \log M_n \leq R.$$

Definition 2.6 (ε -fixed-length source coding rate):

$$L_f(\varepsilon|\mathbf{Y}) = \inf \{R | R \text{ is } \varepsilon\text{-achievable}\}.$$

Then, we have

Theorem 2.3 (Steinberg and Verdú [6]):

$$L_f(\varepsilon|\mathbf{Y}) = \inf \{R | F(R) \leq \varepsilon\} \quad (0 \leq \forall \varepsilon < 1),$$

where $F(R)$ is defined as in (2.1).

The following Fig. 2.3 illustrates Theorem 2.3.

An immediate consequence of Theorem 2.1 and Theorem 2.3 is the following theorem, which reveals an operational equivalence between resolvability and fixed-length source coding from the viewpoint of random number generation, that is,

Theorem 2.4 (Han [5, Remark 2.4.1]):

$$L_f(\varepsilon|\mathbf{Y}) = S_r(2\varepsilon|\mathbf{Y}) \quad (0 \leq \forall \varepsilon < 1).$$

As for the operational meaning of this equivalence, see Appendix A.

III. SECOND-ORDER ASYMPTOTICS

Having reviewed the results on the first-order asymptotics, we now focus on the second-order asymptotics, which will turn out to be in nice correspondence with the first-order asymptotics.

A. Second-Order Resolvability

Definition 3.1: Rate R is said to be (a, δ) -achievable if there exists a mapping $\phi_n : \mathcal{U}_{M_n} \rightarrow \mathcal{Y}^n$ such that

$$\limsup_{n \rightarrow \infty} \frac{1}{\sqrt{n}} \log \frac{M_n}{e^{na}} \leq R \text{ and } \limsup_{n \rightarrow \infty} d(Y^n, \phi_n(U_{M_n})) \leq \delta.$$

Definition 3.2 $((a, \delta)$ -resolvability):

$$S_r(a, \delta | \mathbf{Y}) = \inf \{ R | R \text{ is } (a, \delta)\text{-achievable} \}.$$

Then, we obtain the following fundamental formula for the resolvability problem:

Theorem 3.1:

$$S_r(a, \delta | \mathbf{Y}) = \inf \left\{ R \left| F_a(R) \leq \frac{\delta}{2} \right. \right\} \quad (0 \leq \forall \delta < 2),$$

where

$$F_a(R) = \limsup_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n} \log \frac{1}{P_{Y^n}(Y^n)} \geq a + \frac{R}{\sqrt{n}} \right\}. \quad (3.1)$$

The following Fig. 3.1 illustrates Theorem 3.1.

Remark 3.1: This theorem can be driven as a consequence of Theorem 3.3 combined with Theorem 3.4 below. See also Remark 3.2 in Subsection III-C below.

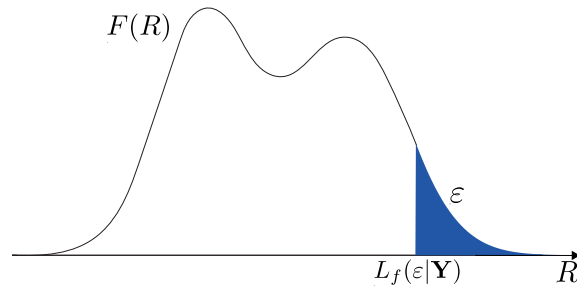


Fig. 2.3. First-Order Fixed-length Source Coding Rate

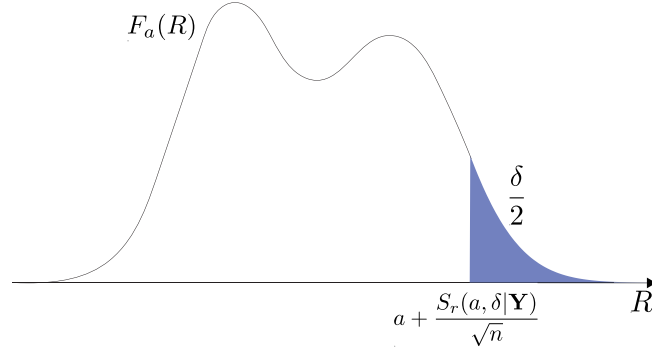


Fig. 3.1. Second-Order Resolvability Rate

B. Second-Order Intrinsic Randomness

Definition 3.3: Rate R is said to be (a, δ) -achievable if there exists a mapping $\phi_n : \mathcal{Y}^n \rightarrow \mathcal{U}_{M_n}$ such that

$$\liminf_{n \rightarrow \infty} \frac{1}{\sqrt{n}} \log \frac{M_n}{e^{na}} \geq R \text{ and } \limsup_{n \rightarrow \infty} d(U_{M_n}, \phi_n(Y^n)) \leq \delta.$$

Definition 3.4 ((a, δ) -intrinsic randomness):

$$S_t(a, \delta | \mathbf{Y}) = \sup \{ R | R \text{ is } (a, \delta)\text{-achievable} \}.$$

Then, we have

Theorem 3.2 (Hayashi [11]):

$$S_t(a, \delta | \mathbf{Y}) = \sup \left\{ R \left| G_a(R) \leq \frac{\delta}{2} \right. \right\} \quad (0 \leq \forall \delta < 2),$$

where

$$G_a(R) = \limsup_{n \rightarrow \infty} \Pr \left\{ \frac{1}{n} \log \frac{1}{P_{Y^n}(Y^n)} \leq a + \frac{R}{\sqrt{n}} \right\}.$$

The following Fig. 3.2 illustrates Theorem 3.2.

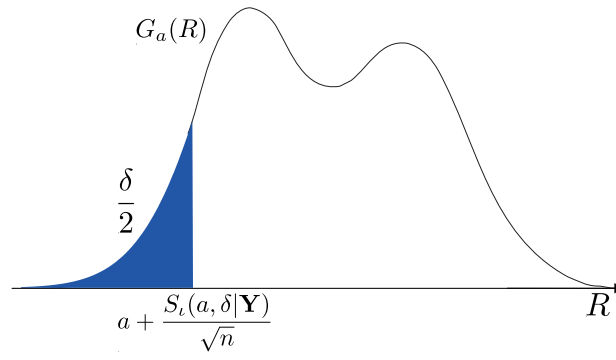


Fig. 3.2. Second-Order Intrinsic Randomness Rate

C. Second-Order Fixed-length Source Coding

Definition 3.5: Rate R is said to be (a, ε) -achievable if there exists a code (n, M_n, ε_n) such that

$$\limsup_{n \rightarrow \infty} \varepsilon_n \leq \varepsilon \text{ and } \limsup_{n \rightarrow \infty} \frac{1}{\sqrt{n}} \log \frac{M_n}{e^{na}} \leq R.$$

Definition 3.6 ((a, ε) -fixed-length source coding rate):

$$L_f(a, \varepsilon | \mathbf{Y}) = \inf \{R | R \text{ is } (a, \varepsilon)\text{-achievable}\}.$$

Then, we have

Theorem 3.3 (Hayashi [11]):

$$L_f(a, \varepsilon | \mathbf{Y}) = \inf \{R | F_a(R) \leq \varepsilon\} \quad (0 \leq \forall \varepsilon < 1)$$

where $F_a(R)$ is defined as in (3.1).

The following Fig. 3.3 illustrates Theorem 3.3.

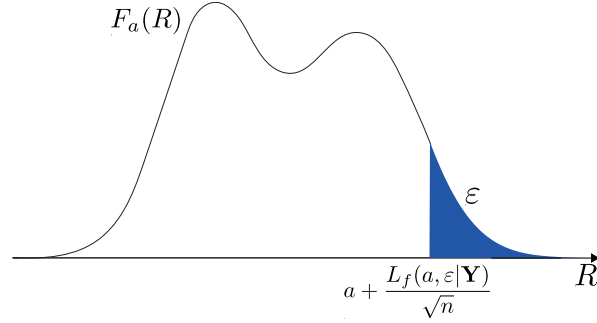


Fig. 3.3. Second-Order Fixed-length Source Coding Rate

On the other hand, we have the following equivalence theorem between resolvability and fixed-length source coding, which is a finer version of Theorem 2.4 for the case of the second-order asymptotics:

Theorem 3.4 (equivalence theorem):

$$L_f(a, \varepsilon | \mathbf{Y}) = S_r(a, 2\varepsilon | \mathbf{Y}) \quad (0 \leq \forall \varepsilon < 1).$$

Proof: This theorem is to reveal a kind of *operational relationship* between resolvability and fixed-length source coding. It suffices to show that if $R > L_f(a, \varepsilon | \mathbf{Y})$ then $R + \gamma > S_r(a, 2\varepsilon | \mathbf{Y})$ for any small $\gamma > 0$, and vice versa with R and $R + \gamma$ swapped. To this end, it is sufficient to literally

follow the arguments described in Han [5, p.163] with γ replaced by $\frac{\gamma}{\sqrt{n}}$. To see its mechanism explicitly, we will give the details of the proof in Appendix A. ■

Remark 3.2: Theorem 3.1 can also directly, not via Theorem 3.4, be proved by using Lemma 5.1 and Lemma 5.2 to be described in Section V. Thus, the second-order resolvability formula for *general* sources can be reasonably established by using the argument similar to that for the first-order formula. This is because the lemmas obtained by the *information spectrum* methods are very basic and fundamental. This viewpoint has been pointed out also in Hayashi [11]. Actually, in [11] the second-order formula for intrinsic randomness (Theorem 3.2) as well as for fixed-length source coding (Theorem 3.3) has been proved using these lemmas (also cf. Lemmas 7.1 and 7.2 in Section VII), which had earlier been used already to establish the first-order formulas for these problems (cf. Han [5]).

IV. SECOND-ORDER ASYMPTOTICS FOR MIXED SOURCES

So far we have demonstrated the general formulas for typical first-order and second-order asymptotic problems (resolvability, intrinsic randomness and fixed-length source coding rate) of random number generation with any general source $\mathbf{Y} = \{Y^n\}_{n=1}^\infty$.

However, computation of these general formulas is quite hard in general or even formidable. Therefore, in this section we consider to introduce a class of tractable sources \mathbf{Y} for which the general formulas are computable but still of independent interest. One of such source classes would be the case where \mathbf{Y} is a *mixed* source of two i.i.d. sources $\mathbf{Y}_1 = \{Y_1^n\}_{n=1}^\infty$ and $\mathbf{Y}_2 = \{Y_2^n\}_{n=1}^\infty$. The computation problem of the first-order asymptotics for such mixed sources has already been solved (e.g., see Han [5]), so in the sequel we now focus on the computation problem of the second-order asymptotics for mixed sources. As a result, it will turn out that we can explicitly compute the asymptotic formulas by virtue of the *information spectrum methods*.

Let us begin with the formal definition of mixed sources. Let $\mathcal{Y} = \{0, 1, 2, \dots\}$ (*countably infinite*) be a source alphabet and $\mathbf{y} = y_1 y_2 \cdots y_n \in \mathcal{Y}^n$ denote a sequence emitted from the source of length n . Let Y^n denote a random variable: a source sequence of length n .

We consider a mixed source consists of two stationary memoryless sources $\mathbf{Y}_i = \{Y_i^n\}_{n=1}^\infty$ with $i = 1, 2$. Then, the mixed source $\mathbf{Y} = \{Y^n\}_{n=1}^\infty$ is defined by

$$P_{Y^n}(\mathbf{y}) = w(1)P_{Y_1^n}(\mathbf{y}) + w(2)P_{Y_2^n}(\mathbf{y}), \quad (4.1)$$

where $w(i)$ are constants satisfying $w(1) + w(2) = 1$ and $w(i) > 0$ ($i = 1, 2$). Since two i.i.d. sources

\mathbf{Y}_i ($i = 1, 2$) are completely specified by giving just the first component Y_i ($i = 1, 2$), we may write simply as $\mathbf{Y}_i = \{Y_i\}$ ($i = 1, 2$) and define the variances:

Definition 4.1 (variance):

$$\sigma_i^2 = E \left(\log \frac{1}{P_{Y_i}(Y_i)} - H(Y_i) \right)^2 \quad (i = 1, 2),$$

where we assume that these variances are finite, and define the entropy by

$$H(Y_i) = \sum_{y \in \mathcal{Y}} P_{Y_i}(y) \log \frac{1}{P_{Y_i}(y)}.$$

Since we consider the case where $\mathbf{Y}_i = \{Y_i\}$ ($i = 1, 2$) is an i.i.d. source, the following asymptotic normality holds for each component i.i.d. source:

$$\lim_{n \rightarrow \infty} \Pr \left\{ \frac{-\log P_{Y_i^n}(Y_i^n) - nH(Y_i)}{\sqrt{n}\sigma_i} \geq U \right\} = \int_U^{+\infty} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{z^2}{2} \right] dz, \quad (4.2)$$

where σ_i^2 denotes the variances defined in Definition 4.1 ($i = 1, 2$).

The following lemma plays the key role in dealing with *mixed* sources in the proof of Theorem 5.1, Theorem 6.1 and Theorem 7.1. In other words, the crux of the arguments for mixed sources in the present paper is summarized by Lemma 4.1, which is completely irrelevant to the matters posed by Hayashi [11].

Lemma 4.1 (Han [5]): Let $\{z_n\}_{n=1}^{\infty}$ be any real-valued sequence. Then for the mixed source \mathbf{Y} it holds that, for $i = 1, 2$,

1)

$$\Pr \left\{ \frac{-\log P_{Y_i^n}(Y_i^n)}{\sqrt{n}} \geq z_n \right\} \geq \Pr \left\{ \frac{-\log P_{Y_i^n}(Y_i^n)}{\sqrt{n}} \geq z_n + \gamma_n \right\} - e^{-\sqrt{n}\gamma_n},$$

2)

$$\Pr \left\{ \frac{-\log P_{Y_i^n}(Y_i^n)}{\sqrt{n}} \geq z_n \right\} \leq \Pr \left\{ \frac{-\log P_{Y_i^n}(Y_i^n)}{\sqrt{n}} \geq z_n - \gamma_n \right\},$$

where $\gamma_n > 0$ satisfies $\gamma_1 > \gamma_2 > \dots > 0$, $\gamma_n \rightarrow 0$, $\sqrt{n}\gamma_n \rightarrow \infty$.

Proof: See Appendix B. ■

V. (a, δ) -RESOLVABILITY

In this section we shall establish $S_r(a, \delta | \mathbf{Y})$ for mixed sources. At first we introduce here two fundamental lemmas of Han [5]: Lemma 5.1 and Lemma 5.2 below. Before describing the lemmas,

we need to define two sets. Let $\mathbf{X} = \{X^n\}_{n=1}^\infty$ and $\mathbf{Y} = \{Y^n\}_{n=1}^\infty$ be arbitrary general sources with *countably infinite* alphabets, and given a sequence $\{z_n\}_{n=1}^\infty$ define $S_n(z_n)$ and $T_n(z_n)$:

$$S_n(z_n) = \left\{ \mathbf{x} \in \mathcal{X}^n \left| \frac{1}{\sqrt{n}} \log \frac{1}{P_{X^n}(\mathbf{x})} \geq z_n \right. \right\},$$

$$T_n(z_n) = \left\{ \mathbf{y} \in \mathcal{Y}^n \left| \frac{1}{\sqrt{n}} \log \frac{1}{P_{Y^n}(\mathbf{y})} \leq z_n \right. \right\}.$$

Lemma 5.1: Let $\mathbf{X} = \{X^n\}_{n=1}^\infty$ and $\mathbf{Y} = \{Y^n\}_{n=1}^\infty$ be arbitrary general sources, where X^n and Y^n are random variables taking values in \mathcal{X}^n and \mathcal{Y}^n , respectively. Then, for an arbitrary sequence $\{z_n\}_{n=1}^\infty$ and $\gamma > 0$, there exists a mapping $\phi_n : \mathcal{X}^n \rightarrow \mathcal{Y}^n$ such that

$$d(\phi_n(X^n), Y^n) \leq 2 \max(\Pr\{X^n \notin S_n(z_n + \gamma)\}, \Pr\{Y^n \notin T_n(z_n)\}) + 2e^{-\sqrt{n}\gamma}.$$

Lemma 5.2: Let $\mathbf{X} = \{X^n\}_{n=1}^\infty$ and $\mathbf{Y} = \{Y^n\}_{n=1}^\infty$ be arbitrary general sources, where X^n and Y^n are random variables taking values in \mathcal{X}^n and \mathcal{Y}^n , respectively. Then, for an arbitrary sequence $\{z_n\}_{n=1}^\infty$, $\gamma > 0$ and any mapping $\phi_n : \mathcal{X}^n \rightarrow \mathcal{Y}^n$ it holds that

$$d(\phi_n(X^n), Y^n) \geq 2 \Pr\{Y^n \notin T_n(z_n + \gamma)\} - 2 \Pr\{X^n \in S_n(z_n)\} - 2e^{-\sqrt{n}\gamma}. \quad (5.1)$$

Remark 5.1: Also, (5.1) can be written as

$$d(\phi_n(X^n), Y^n) \geq 2 \Pr\{X^n \notin S_n(z_n)\} - 2 \Pr\{Y^n \in T_n(z_n + \gamma)\} - 2e^{-\sqrt{n}\gamma}. \quad (5.2)$$

The above lemmas are useful for the random number generation problem to approximate a probability distribution $\mathbf{Y} = \{Y^n\}_{n=1}^\infty$ by using an another probability distribution $\mathbf{X} = \{X^n\}_{n=1}^\infty$. Clearly, this includes the resolvability problem as a special case, which is the case of $X^n = U_{M_n}$. Therefore, in this case the condition in the above lemmas leads to

$$\Pr\{X^n \notin S_n(z_n)\} = \begin{cases} 0 & z_n \leq \frac{1}{\sqrt{n}} \log M_n \\ 1 & z_n > \frac{1}{\sqrt{n}} \log M_n. \end{cases} \quad (5.3)$$

Notice that the above lemmas are valid for general sources \mathbf{X} and \mathbf{Y} .

In the sequel, we consider the case that $0 \leq \delta < 2$ and $w(1) \neq \frac{\delta}{2}$ hold (cf. Remark 5.2 below for the case of $w(1) = \frac{\delta}{2}$). Then, given $0 \leq \delta < 2$ we classify the problem into three cases. Here, without loss of generality, we assume that $H(Y_1) \geq H(Y_2)$ holds:

- I $H(Y_1) = H(Y_2)$ holds.
- II $H(Y_1) > H(Y_2)$ and $w(1) > \frac{\delta}{2}$ hold.
- III $H(Y_1) > H(Y_2)$ and $w(1) < \frac{\delta}{2}$ hold.

In Case I, we shall establish $S_r(H(Y_1), \delta | \mathbf{Y})$ (obviously, this is equal to $S_r(H(Y_2), \delta | \mathbf{Y})$). In Case II and Case III we shall show $S_r(H(Y_1), \delta | \mathbf{Y})$ and $S_r(H(Y_2), \delta | \mathbf{Y})$, respectively. Now we have one of the main results:

Theorem 5.1: Given $0 \leq \delta < 2$, the following holds.

Case I:

$$S_r(H(Y_1), \delta | \mathbf{Y}) = T_1, \quad (5.4)$$

where T_1 is specified by

$$\frac{\delta}{2} = \sum_{i=1}^2 w(i) \int_{\frac{T_1}{\sigma_i}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{z^2}{2} \right] dz. \quad (5.5)$$

Case II:

$$S_r(H(Y_1), \delta | \mathbf{Y}) = T_2, \quad (5.6)$$

where T_2 is specified by

$$\frac{\delta}{2} = w(1) \int_{\frac{T_2}{\sigma_1}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{z^2}{2} \right] dz. \quad (5.7)$$

Case III:

$$S_r(H(Y_2), \delta | \mathbf{Y}) = T_3, \quad (5.8)$$

where T_3 is specified by

$$\frac{\delta}{2} = w(1) + w(2) \int_{\frac{T_3}{\sigma_2}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{z^2}{2} \right] dz. \quad (5.9)$$

Illustrative figures in each case of this theorem are depicted in Fig. 5.1–Fig. 5.3, where the weighted probability of the shaded area is equal to $\frac{\delta}{2}$.

Remark 5.2: It is easy to check that $T_2 = -\infty$, $T_3 = +\infty$ for $w(1) = \frac{\delta}{2}$. Also, it will turn out from the way of proving the above theorem that the second-order asymptotics gets trivial if $a \neq H(Y_1)$ and $a \neq H(Y_2)$, because this case necessarily implies that $\delta = 0$ or $\delta = 2w(1)$ or $\delta = 2$, depending on the value of a ; then, accordingly, we can formally set as $S_r(a, \delta | \mathbf{Y}) = -\infty$.

Remark 5.3: Theorem 5.1 can be restated more intuitively but equivalently as follows. Let

$$R_n \equiv \frac{1}{n} \log M_n$$

denote the *size rate* of resolvability (cf. Definition 3.1 and 3.2), and consider the following asymptotic equation for R_n :

$$w(1)\Phi\left(\frac{\sqrt{n}(R_n - H(Y_1))}{\sigma_1}\right) + w(2)\Phi\left(\frac{\sqrt{n}(R_n - H(Y_2))}{\sigma_2}\right) = \frac{\delta}{2}, \quad (5.10)$$

where $\Phi(\cdot)$ is the Gaussian cumulative distribution function defined by

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} e^{-\frac{x^2}{2}} dx.$$

Denote the solution of this equation by

$$R_n^* = \frac{1}{n} \log M_n^*$$

and set as

$$R_n^* = a + \frac{b}{\sqrt{n}} + o\left(\frac{1}{\sqrt{n}}\right) \quad (a, b \text{ are constants}),$$

which, substituted into (5.10), yields

$$w(1)\Phi\left(\frac{\sqrt{n}(a - H(Y_1)) + b + o(1)}{\sigma_1}\right) + w(2)\Phi\left(\frac{\sqrt{n}(a - H(Y_2)) + b + o(1)}{\sigma_2}\right) = \frac{\delta}{2}. \quad (5.11)$$

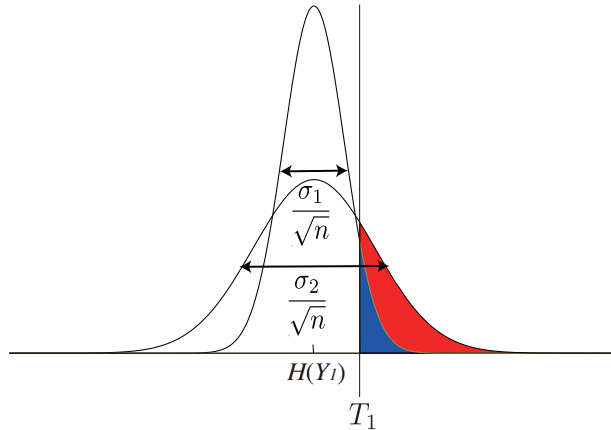


Fig. 5.1. Case I

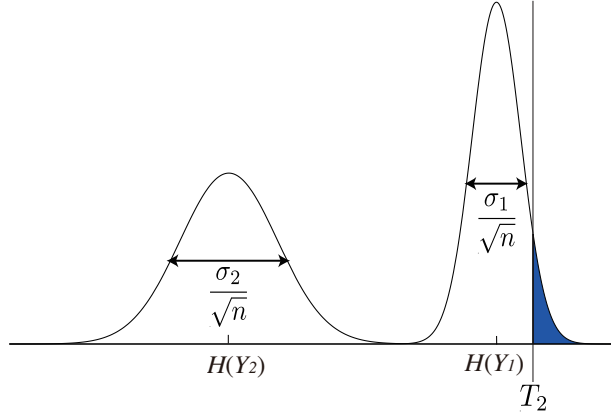


Fig. 5.2. Case II

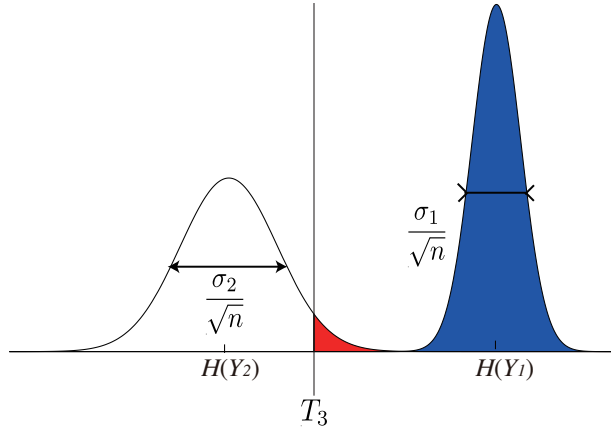


Fig. 5.3. Case III

Then, it is not difficult to verify by letting $n \rightarrow \infty$ that, given a and δ , the corresponding solution $b = b^*(a, \delta)$ of this equation coincides with T_1 , T_2 and T_3 , respectively, according to Cases I, II, III (cf. also (8.3) in the proof of Theorem 8.1). Notice here that the equation (5.10) subsumes Remark 5.2 too. Thus, it is concluded that $b^*(a, \delta)$ is nothing but the second-order resolvability $S_r(a, \delta | \mathbf{Y})$, and hence Theorem 5.1 is equivalent to the equation (5.10).

Summarizing up, we can write the optimal size M_n^* as

$$\log M_n^* = na + \sqrt{n}b^*(a, \delta) + o(\sqrt{n}),$$

with $a = H(Y_1)$ or $a = H(Y_2)$ depending on the value of δ , that is, if $H(Y_1) \neq H(Y_2)$ then

$$\begin{aligned} a = H(Y_1) \text{ and } b^*(a, \delta) &= \sigma_1 \Phi^{-1} \left(\frac{\delta}{2w(1)} \right) & \text{if } w(1) > \frac{\delta}{2}; \\ a = H(Y_2) \text{ and } b^*(a, \delta) &= \sigma_2 \Phi^{-1} \left(\frac{\delta - 2w(1)}{2w(2)} \right) & \text{if } w(1) < \frac{\delta}{2}, \end{aligned} \quad (5.12)$$

which enables us to evaluate how large size of M_n^* is needed as a function of block length n and variational distance δ . Notice here that in this case $b^*(a, \delta)$ can be written as the simple inverse function Φ^{-1} of the Gaussian distribution function, and also that $b^*(a, \delta)$ can be negative, for example, $b^*(a, \delta) < 0$ if $\delta > 1 + w(1)$, so that in this case the first-order δ -resolvability $S_r(\delta|\mathbf{Y})$ is $H(Y_2)$ but the optimal achievable rate R_n^* approaches it *from below*. In other words, it is possible to make necessary rates to be below the δ -resolvability at finite block length n . The nonergodic channel counterpart of the equations (5.10) and (5.12) has been provided by Polyanskiy, Poor and Verdú [13], who have observed for the Gilbert Elliott channel the same kind of non-asymptotic phenomenon as here. On the other hand, in the case where $H(Y_1) = H(Y_2)$ holds, $b^*(a, \delta)$ can be written as the inverse function of a mixed Gaussian distribution function (see, Case I). Notice here that Case I is missing in [13]. ■

Proof of Theorem 5.1: See Appendix C. ■

VI. (a, δ) -INTRINSIC RANDOMNESS

Let us now turn to the computation problem of the (a, δ) -intrinsic randomness formula for mixed sources. To do so, without loss of generality, we consider the following three cases:

- I $H(Y_1) = H(Y_2)$ holds.
- II $H(Y_1) > H(Y_2)$ and $w(2) > \frac{\delta}{2}$ hold.
- III $H(Y_1) > H(Y_2)$ and $w(2) < \frac{\delta}{2}$ hold.

Then, we have

Theorem 6.1: Given $0 \leq \forall \delta < 2$, the following holds.

Case I:

$$S_\iota(H(Y_2), \delta|\mathbf{Y}) = T_4,$$

where T_4 is specified by

$$\frac{\delta}{2} = \sum_{i=1}^2 w(i) \int_{-\infty}^{\frac{T_4}{\sigma_i}} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{z^2}{2} \right] dz.$$

Case II:

$$S_\ell(H(Y_2), \delta|\mathbf{Y}) = T_5,$$

where T_5 is specified by

$$\frac{\delta}{2} = w(2) \int_{-\infty}^{\frac{T_5}{\sigma_2}} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{z^2}{2}\right] dz.$$

Case III:

$$S_\ell(H(Y_1), \delta|\mathbf{Y}) = T_6,$$

where T_6 is specified by

$$\frac{\delta}{2} = w(2) + w(1) \int_{-\infty}^{\frac{T_6}{\sigma_1}} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{z^2}{2}\right] dz.$$

Proof: It suffices to proceed in parallel with the arguments as made in the proof of Theorem 5.1, while taking account of the *duality* between resolvability and intrinsic randomness. Note that (5.2) is used instead of (5.1) in the proof of Converse Part. \blacksquare

Remark 6.1: It is easy to see that $T_5 = +\infty$ and $T_6 = -\infty$ for $w(2) = \frac{\delta}{2}$. The latter part of Remark 5.2 similarly applies here too with $S_\ell(a, \delta|\mathbf{Y}) = +\infty$ instead of $S_r(a, \delta|\mathbf{Y}) = -\infty$. Also, like in Remark 5.3, we can ascertain that Theorem 6.1 is equivalent to the asymptotic equation:

$$w(1)\Psi\left(\frac{\sqrt{n}(R_n - H(Y_1))}{\sigma_1}\right) + w(2)\Psi\left(\frac{\sqrt{n}(R_n - H(Y_2))}{\sigma_2}\right) = \frac{\delta}{2},$$

where $\Psi(x) = 1 - \Phi(x)$ and T_4, T_5, T_6 instead of T_1, T_2, T_3 , respectively.

VII. (a, ε) -FIXED-LENGTH SOURCE CODING

Let us now consider to compute the formula for $L_f(a, \varepsilon|\mathbf{Y})$ for mixed sources. To do so, without loss of generality, we consider the following three cases:

- I $H(Y_1) = H(Y_2)$ holds.
- II $H(Y_1) > H(Y_2)$ and $w(1) > \varepsilon$ hold.
- III $H(Y_1) > H(Y_2)$ and $w(1) < \varepsilon$ hold.

Then, we have the following main result:

Theorem 7.1: Given $0 \leq \varepsilon < 1$, the following holds.

Case I:

$$L_f(H(Y_1), \varepsilon|\mathbf{Y}) = T_7, \tag{7.1}$$

where T_7 is specified by

$$\varepsilon = \sum_{i=1}^2 w(i) \int_{\frac{T_7}{\sigma_i}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{z^2}{2} \right] dz. \quad (7.2)$$

Case II:

$$L_f(H(Y_1), \varepsilon | \mathbf{Y}) = T_8, \quad (7.3)$$

where T_8 is specified by

$$\varepsilon = w(1) \int_{\frac{T_8}{\sigma_1}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{z^2}{2} \right] dz.$$

Case III:

$$L_f(H(Y_2), \varepsilon | \mathbf{Y}) = T_9,$$

where T_9 is specified by

$$\varepsilon = w(1) + w(2) \int_{\frac{T_9}{\sigma_2}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{z^2}{2} \right] dz.$$

Remark 7.1: It is easy to check that $T_8 = -\infty$, $T_9 = +\infty$ for $w(1) = \varepsilon$. The latter part of Remark 5.2 similarly applies here too with ε instead of $\frac{\delta}{2}$ and $L_f(a, \varepsilon | \mathbf{Y}) = -\infty$ instead of $S_r(a, \delta | \mathbf{Y}) = -\infty$. Remark 5.3 applies also here with ε instead of $\frac{\delta}{2}$, and T_7 , T_8 , T_9 instead of T_1 , T_2 , T_3 , respectively. In particular, it turns out that Theorem 7.1 is equivalent to the equation (5.10).

Proof of Theorem 7.1: Although the proof is immediate from Theorem 3.4 and Theorem 5.1 with $\varepsilon = \frac{\delta}{2}$, we give in Appendix D another information spectrum approach that is of independent interest, where Lemma 7.1 and Lemma 7.2 (due to Han [5]) as described below are invoked, which Hayashi [11] has fully used to obtain the *general* formula for the second-order optimal fixed-length source coding rate (see, Theorem 3.3) but not for *mixed* sources. ■

Lemma 7.1: Let M_n be an arbitrary given positive integer. Then, for all $n = 1, 2, \dots$, there exists an (n, M_n, ε_n) code such that

$$\varepsilon_n \leq \Pr \left\{ \frac{1}{\sqrt{n}} \log \frac{1}{P_{Y^n}(Y^n)} \geq \frac{1}{\sqrt{n}} \log M_n \right\}.$$

Lemma 7.2: For all $n = 1, 2, \dots$, any (n, M_n, ε_n) code satisfies

$$\varepsilon_n \geq \Pr \left\{ \frac{1}{\sqrt{n}} \log \frac{1}{P_{Y^n}(Y^n)} \geq \frac{1}{\sqrt{n}} \log M_n + \gamma \right\} - e^{-\sqrt{n}\gamma},$$

where $\gamma > 0$ is an arbitrary constant.

VIII. SECOND-ORDER RESOLVABILITY FOR MIXED SOURCES: GENERAL CASES

In the previous sections, we have established the second-order achievable rates for a mixture of two i.i.d. sources, which is regarded as being the simplest model of mixed sources. In this section, we shall extend our results to more general mixed sources. Although we consider only the resolvability problem in this section, similar extensions can be done immediately for the intrinsic randomness problem and the fixed-length coding problem.

A. (a, δ) -Resolvability for a Mixture of Countably Infinite i.i.d. Sources

Let us now consider the mixed source consisting of *countably infinite* stationary memoryless sources $\mathbf{Y}_i = \{Y_i^n\}_{n=1}^\infty$, where $i \in \mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$. The mixed source that we consider in this subsection is defined by

$$P_{Y^n}(\mathbf{y}) = \sum_{i=-\infty}^{\infty} w(i) P_{Y_i^n}(\mathbf{y}),$$

where $w(i)$ are constants such that $\sum_{i=-\infty}^{\infty} w(i) = 1$ and $w(i) \geq 0$ for all $i \in \mathbb{Z}$. The variance σ_i^2 ($i \in \mathbb{Z}$) of the stationary memoryless source \mathbf{Y}_i is defined in a similar way to Definition 4.1 and we assume that all these variances are finite. Finally, without loss of generality we assume that

$$\dots \leq H(Y_{-2}) \leq H(Y_{-1}) \leq H(Y_0) \leq H(Y_1) \leq H(Y_2) \leq \dots$$

With this definition of mixed sources, we now have the following second-order resolvability theorem, which is a substantial generalization of Theorem 5.1.

Theorem 8.1: For a mixture \mathbf{Y} of *countably infinite* i.i.d. sources with *countably infinite* alphabet, the optimal size rate $R_n^* = \frac{1}{n} \log M_n^*$ is given as the solution for R_n of the asymptotic equation:

$$\sum_{i=-\infty}^{\infty} w(i) \Phi \left(\frac{\sqrt{n}(R_n - H(Y_i))}{\sigma_i} \right) = \frac{\delta}{2}; \quad (8.1)$$

and, furthermore, the resolvability $S_r(a, \delta | \mathbf{Y})$ is given as the solution $b = b^*(a, \delta)$ of the asymptotic equation

$$\sum_{i=-\infty}^{\infty} w(i) \Phi \left(\frac{\sqrt{n}(a - H(Y_i)) + b + o(1)}{\sigma_i} \right) = \frac{\delta}{2}. \quad (8.2)$$

Proof: The derivation of equation (8.1) and (8.2) is based on the *multi-peak* asymptotic normality instead of the two-peak asymptotic normality. Although it is easy to verify that Theorem 8.1 follows in a manner similar to that in Remark 5.3, for the sake of reader's convenience we demonstrate here also the formula of the form as in Theorem 5.1. Define the sets of indices as

$$\mathcal{L}_0(a) \equiv \{k \in \mathbb{Z} | H(Y_k) = a, w(k) > 0\}.$$

$$\mathcal{L}_1(a) \equiv \{i \in \mathbb{Z} | H(Y_i) > a, w(i) > 0\},$$

We first consider the case where $\mathcal{L}_0(a) = \emptyset$ (the empty set). This case necessarily implies

$$\sum_{i \in \mathcal{L}_1(a)} w(i) = \frac{\delta}{2},$$

which allows us to formally set as $S_r(a, \delta | \mathbf{Y}) = -\infty$.

We next consider the case of $\mathcal{L}_0(a) \neq \emptyset$. It is easy to see that letting $n \rightarrow \infty$ in (8.2) yields the following *non-asymptotic* equation to determine the second-order resolvability $b = b^*(a, \delta)$:

$$\sum_{i \in \mathcal{L}_0(a)} w(i) \Phi\left(\frac{b}{\sigma_i}\right) + \sum_{j \in \mathcal{L}_1(a)} w(j) = \frac{\delta}{2}. \quad (8.3)$$

In order to specifically compute the $b = b^*(a, \delta)$, we need several classifications as in the proof of Theorem 5.1, where equation (8.3) implies that a given δ is specified by the following conditions:

If $\mathcal{L}_0(a) = \emptyset$ then $\sum_{j \in \mathcal{L}_1(a)} w(j) = \frac{\delta}{2}$ ($b = -\infty$); If $\mathcal{L}_0(a) \neq \emptyset$ then

$$\begin{aligned} \sum_{j \in \mathcal{L}_1(a)} w(j) &< \frac{\delta}{2}, \\ \sum_{i \in \mathcal{L}_0(a)} w(i) + \sum_{j \in \mathcal{L}_1(a)} w(j) &\geq \frac{\delta}{2}. \end{aligned}$$

This is illustrated in Fig. 8.1.

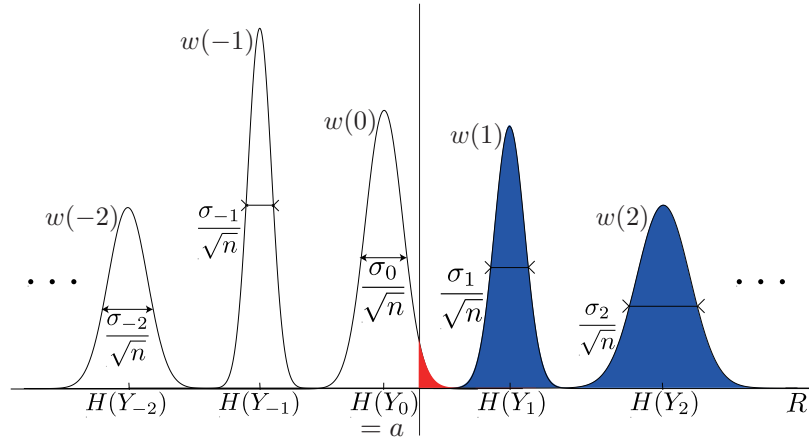


Fig. 8.1. Countably Infinite i.i.d. Sources

It is easy to see that these conditions yield Cases I, II, III in Theorem 5.1 as a special case. ■

B. (a, δ) -Resolvability for a Mixture of Countably Infinite Markovian Sources

Next, let us consider a mixed source consisting of *countably infinite* Markovian sources with *finite* source alphabet $\mathbf{Y}_i = \{Y_i^n\}_{n=1}^\infty$, where $i \in \mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$. The mixed source that we consider in this subsection is defined by

$$P_{Y^n}(\mathbf{y}) = \sum_{i=-\infty}^{\infty} w(i) P_{Y_i^n}(\mathbf{y}),$$

where $w(i)$ are constants satisfying $\sum_{i=-\infty}^{\infty} w(i) = 1$ and $w(i) \geq 0$ for all $i \in \mathbb{Z}$ and each $\mathbf{Y}_i = \{Y_i^n\}_{n=1}^\infty$ denotes the stationary ergodic Markovian source with irreducible transition probability $Q_i(k|j)$ ($(j, k) \in \mathcal{Y}^2$), where j and k denote consecutive symbols in the Markov transition. We let $\pi_i(\cdot)$ denote the stationary distribution of \mathbf{Y}_i . The probability of a sequence $\mathbf{y} = y_1 y_2 \dots y_n \in \mathcal{Y}^n$ emitted from Markovian source \mathbf{Y}_i is given by

$$P_{Y_i^n}(\mathbf{y}) = \pi_i(y_1) \prod_{j=1}^{n-1} Q_i(y_{j+1}|y_j).$$

The entropy rate $H(Q_i)$ ($i \in \mathbb{Z}$) and the variance σ_i^2 ($i \in \mathbb{Z}$) of Markovian source \mathbf{Y}_i are defined by

$$H(Q_i) \equiv \sum_{(k,j) \in \mathcal{Y}^2} \pi_i(j) Q_i(k|j) \log \frac{1}{Q_i(k|j)},$$

and

$$\begin{aligned} \sigma_i^2 \equiv & \sum_{(k,j) \in \mathcal{Y}^2} \pi_i(j) Q_i(k|j) \left(\log \frac{1}{Q_i(k|j)} - H(Q_i) \right)^2 \\ & + 2 \sum_{(l,k,j) \in \mathcal{Y}^3} \pi_i(j) Q_i(l|k) Q_i(k|j) \left(\log \frac{1}{Q_i(l|k)} - H(Q_i) \right) \left(\log \frac{1}{Q_i(k|j)} - H(Q_i) \right), \end{aligned}$$

respectively (see Hayashi [11, VII]). In addition, we assume that all these variances are finite, and it is also assumed that

$$\dots \leq H(Q_{-2}) \leq H(Q_{-1}) \leq H(Q_0) \leq H(Q_1) \leq H(Q_2) \leq \dots$$

With this definition of mixed sources, we now have the following second-order resolvability theorem:

Theorem 8.2: For a mixture \mathbf{Y} of *countably infinite* Markovian sources \mathbf{Y}_i ($i \in \mathbb{Z}$) with *finite* source alphabet, the optimal size rate $R_n^* = \frac{1}{n} \log M_n^*$ is given as the solution of the asymptotic equation:

$$\sum_{i=-\infty}^{\infty} w(i) \Phi \left(\frac{\sqrt{n}(R_n - H(Q_i))}{\sigma_i} \right) = \frac{\delta}{2};$$

and, furthermore, the second-order resolvability $S_r(a, \delta | \mathbf{Y})$ is given as the solution $b = b^*(a, \delta)$ of the equation

$$\sum_{i=-\infty}^{\infty} w(i) \Phi \left(\frac{\sqrt{n}(a - H(Q_i)) + b + o(1)}{\sigma_i} \right) = \frac{\delta}{2}.$$

Proof: Since we assume that the source alphabet \mathcal{Y} is finite, the number of states of this Markov chain is finite. Then, the asymptotic normality also holds due to the central limit theorem for Markov chains (see, e.g., Chung [14], Billingsley [15]). Therefore, it suffices to proceed in parallel with the arguments as made in the proof of Theorem 8.1 and so we omit the details. ■

Remark 8.1: As shown in this section, our analysis for mixed sources is on the basis of the asymptotic normality of self-information. This means that the similar argument is valid for any mixture of countably infinite sources for which the asymptotic normality of self-information holds for each of the component sources. Generally speaking, one of the conditions to ensure the central limit theorem for weakly dependent random variables is, e.g., the *mixing* condition (see, Billingsley [15], Shields [16]).

C. (a, δ) -Resolvability for Sources with General Mixture

In this subsection we consider a possible extension of Theorems 8.1 and 8.2 to the case with *general mixture* instead of countably infinite mixture. It suffices here only to focus on the extension of Theorem 8.1, because the basic logics underlying both extensions are the same. A mixed source $\mathbf{Y} = \{Y^n\}_{n=1}^{\infty}$ with *general mixture* is defined by

$$P_{Y^n}(\mathbf{y}) = \int_{\Lambda} P_{Y_{\theta}^n}(\mathbf{y}) dw(\theta), \quad (8.4)$$

where $w(\theta)$ is an arbitrary probability measure on the parameter space Λ , and $\mathbf{Y}_{\theta} = \{Y_{\theta}^n\}_{n=1}^{\infty}$ ($\theta \in \Lambda$) are i.i.d. sources with finite alphabet.

With this definition, we have the following formal extension of Theorem 8.1:

Theorem 8.3: For a source \mathbf{Y} with general mixture $w(\theta)$ of i.i.d. sources \mathbf{Y}_{θ} with finite alphabet, the optimal size rate $R_n^* = \frac{1}{n} \log M_n^*$ is given as the solution for R_n of the asymptotic equation:

$$\int_{\Lambda} \Phi \left(\frac{\sqrt{n}(R_n - H(Y_{\theta}))}{\sigma_{\theta}} \right) dw(\theta) = \frac{\delta}{2};$$

and, furthermore, the second-order resolvability $S_r(a, \delta | \mathbf{Y})$ is given as the solution $b = b^*(a, \delta)$ of the asymptotic equation

$$\int_{\Lambda} \Phi \left(\frac{\sqrt{n}(a - H(Y_{\theta})) + b + o(1)}{\sigma_{\theta}} \right) dw(\theta) = \frac{\delta}{2},$$

where σ_θ^2 is the variance of $\log \frac{1}{P_{Y_\theta}(Y_\theta)}$.

Proof: The second inequality of Lemma 4.1 is not necessarily valid in the case with general mixture. Nevertheless, we can slightly modify it so as to be applicable to this general case. As for the details, see the proof of Han [5, Lemma 1.4.4]. ■

Remark 8.2: Let us define, instead of $\mathcal{L}_0(a)$ and $\mathcal{L}_1(a)$ as above, the subsets $\Lambda_0(a)$ and $\Lambda_1(a)$ of Λ as follows:

$$\Lambda_0(a) = \{\theta \in \Lambda | H(Y_\theta) = a\},$$

$$\Lambda_1(a) = \{\theta \in \Lambda | H(Y_\theta) > a\}.$$

Then, the equation (8.3) to determine the (a, δ) -resolvability $b = b^*(a, \delta)$ is extended accordingly to the case with general mixture $w(\theta)$ as follows:

$$\int_{\Lambda_0(a)} \Phi\left(\frac{b}{\sigma_\theta}\right) dw(\theta) + \int_{\Lambda_1(a)} dw(\theta) = \frac{\delta}{2}, \quad (8.5)$$

which implies that a given δ is specified by the following conditions: If $\int_{\Lambda_0(a)} dw(\theta) = 0$ then $\int_{\Lambda_1(a)} dw(\theta) = \frac{\delta}{2}$ ($b = -\infty$); If $\int_{\Lambda_0(a)} dw(\theta) > 0$ then

$$\begin{aligned} \int_{\Lambda_1(a)} dw(\theta) &< \frac{\delta}{2}, \\ \int_{\Lambda_0(a)} dw(\theta) + \int_{\Lambda_1(a)} dw(\theta) &\geq \frac{\delta}{2}. \end{aligned}$$

IX. CONCLUDING REMARKS

We have so far considered the second-order achievability to evaluate the finer structure of random number generation for mixed sources. The class of mixed sources is very important, because all of stationary sources can be regarded as forming mixed sources obtained by mixing stationary ergodic sources with respect to appropriate probability measures. Although, in general, mixed sources do not have the asymptotic normality. So, our result is also meaningful, we have demonstrated that the analysis based on the *two-peak* asymptotic normality is still effective also for sources whose self-information spectrum does *not* have a single asymptotic normality.

As shown in the proofs of the present paper, the information spectrum approach is substantial in the analysis of the *second-order* achievable rates. In particular, Lemma 4.1 is a simple but enables us to work with the *multi-peak* asymptotic normality for mixed sources (cf. the proof of Theorem 8.1).

Polyanskiy, Poor and Verdú [13] has derived the *second-order* capacity in either *ergodic* or *nonergodic* setting for the Gilbert-Elliott channel (GEC), which consists of two *binary* symmetric channels. Their results for the *nonergodic* case is also based on a mixture of two Gaussian distributions. On the other hand, since our analysis is based on the information spectrum methods, the results are valid also for a mixture of *countably infinite* sources (i.i.d. with *countably infinite* alphabet or Markov with *finite* alphabet) as well as a *general mixture* of i.i.d. sources with *finite* alphabet as was shown in Section VIII. It should be emphasized that, throughout in the paper, we have established the *asymptotic* or *nonasymptotic* equations for determining the second-order resolvability, intrinsic randomness, and fixed-length source coding rate. The forms of these equations include that as shown in [13] as a special case. We observe that the second-order capacity of finite state Markov channels (see, [17]) or Fritchman channels (see, [18]), which is generalizations of the GEC, can be established by means of the similar arguments. This will be reported in a forthcoming paper [19].

APPENDIX A

PROOF OF THEOREM 3.4

This appendix concerns the operational relationship between resolvability and fixed-length source coding. Although the proof literally mimics the argument given by Han [5, p.163], we will repeat it here for the reader's convenience.

First, we show how to construct an (n, M_n, ε_n) source code, given a resolvability mapping $\phi_n : \mathcal{U}_{M_n} \rightarrow \mathcal{Y}^n$. Set $\tilde{Y}^n = \phi_n(U_{M_n})$, then we can define the subset \mathcal{S}_0 of \mathcal{Y}^n by

$$\mathcal{S}_0 = \{\mathbf{y} \in \mathcal{Y}^n \mid P_{\tilde{Y}^n}(\mathbf{y}) > 0\}.$$

Clearly, we have $|\mathcal{S}_0| \leq M_n$ which enables us to define the source encoder $\varphi_n : \mathcal{Y}^n \rightarrow \mathcal{U}_{M_n}$ which transforms each element of \mathcal{S}_0 to a distinct element of \mathcal{U}_{M_n} and all elements of $\mathcal{Y}^n \setminus \mathcal{S}_0$ to 1. In addition, if we define the source decoder ψ_n as the inverse mapping of $\varphi_n|_{\mathcal{S}_0}$, the (φ_n, ψ_n) becomes an (n, M_n, ε_n) source code for the source $\mathbf{Y} = \{Y^n\}_{n=1}^\infty$ such that

$$\varepsilon_n = \Pr\{Y^n \notin \mathcal{S}_0\}.$$

Here, from the other definition of the variational distance:

$$d(Z, \tilde{Z}) = 2 \sup_{A: A \subseteq \mathcal{Z}} |P_Z(A) - P_{\tilde{Z}}(A)|,$$

it follows that

$$\Pr\{\tilde{Y}^n \in \mathcal{S}_0\} - \Pr\{Y^n \in \mathcal{S}_0\} \leq \frac{1}{2}d(Y^n, \tilde{Y}^n).$$

Considering that $\Pr\{\tilde{Y}^n \in \mathcal{S}_0\} = 1$ holds, we see that

$$\Pr\{Y^n \notin \mathcal{S}_0\} \leq \frac{1}{2}d(Y^n, \tilde{Y}^n).$$

Consequently, we obtain

$$\varepsilon_n \leq \frac{1}{2}d(Y^n, \tilde{Y}^n). \quad (\text{A.1})$$

In this way we can construct an (n, M_n, ε_n) source code satisfying (A.1) with the same M_n as in the resolvability mapping $\phi_n : \mathcal{U}_{M_n} \rightarrow \mathcal{Y}^n$, which implies that if $R > S_r(a, \delta|\mathbf{Y})$ then $R > L_f(a, \frac{\delta}{2}|\mathbf{Y})$.

Next, let us show how to construct a resolvability mapping ϕ_n , given an (n, M_n, ε_n) source code. We can define the subset \mathcal{S}_0 of \mathcal{Y}^n by

$$\mathcal{S}_0 = \{\mathbf{y} \in \mathcal{Y}^n \mid \mathbf{y} = \psi_n(\varphi_n(\mathbf{y}))\},$$

where (φ_n, ψ_n) denotes the pair of encoder and decoder of the (n, M_n, ε_n) code. For an arbitrarily small $\gamma > 0$ set $M'_n = M_n e^{\sqrt{n}\gamma}$. From the same argument as in the proof of Lemma 5.1 (cf. Han [5]) with $U_{M'_n}$, \mathcal{S}_0 and $\mathcal{U}_{M'_n}$ in place of X^n , $T_n(z_n)$ and $S_n(z_n + \gamma)$, we can construct a mapping $\phi_n : \mathcal{U}_{M'_n} \rightarrow \mathcal{Y}^n$ such that

$$d(Y^n, \tilde{Y}^n) \leq 2\Pr\{Y^n \notin \mathcal{S}_0\} + 2e^{-\sqrt{n}\gamma}, \quad (\text{A.2})$$

where $\tilde{Y}^n = \phi_n(U_{M'_n})$. Since $\varepsilon_n = \Pr\{Y^n \notin \mathcal{S}_0\}$, (A.2) can be expressed as

$$d(Y^n, \tilde{Y}^n) \leq 2\varepsilon_n + 2e^{-\sqrt{n}\gamma}.$$

In this way we can construct the resolvability mapping $\tilde{Y}^n = \phi_n(U_{M'_n})$, which implies that if $R > L_f(a, \varepsilon|\mathbf{Y})$ then $R + \gamma > S_r(a, 2\varepsilon|\mathbf{Y})$.

APPENDIX B

PROOF OF LEMMA 4.1

Although the proof to be shown below is implicitly contained in Han [5], we explicitly summarize it here for the reader's convenience. At first we show the first inequality. Set a sequence $\{\gamma_n\}_{n=1}^\infty$ satisfying $\gamma_1 > \gamma_2 > \dots > 0$, $\gamma_n \rightarrow 0$, $\sqrt{n}\gamma_n \rightarrow \infty$. Then it holds that

$$\begin{aligned} & \Pr\left\{\frac{-\log P_{Y^n}(Y_i^n)}{\sqrt{n}} - \frac{-\log P_{Y_i^n}(Y_i^n)}{\sqrt{n}} \leq -\gamma_n\right\} \\ &= \sum_{\mathbf{y} \in D_n(i)} P_{Y_i^n}(\mathbf{y}) \\ &\leq \sum_{\mathbf{y} \in D_n(i)} P_{Y^n}(\mathbf{y}) e^{-\sqrt{n}\gamma_n} \leq e^{-\sqrt{n}\gamma_n}, \end{aligned}$$

for $i = 1, 2$, where

$$D_n(i) = \left\{ \mathbf{y} \in \mathcal{Y}^n \left| \frac{-\log P_{Y^n}(\mathbf{y})}{\sqrt{n}} - \frac{-\log P_{Y_i^n}(\mathbf{y})}{\sqrt{n}} \leq -\gamma_n \right. \right\}.$$

This means that

$$\Pr \left\{ \frac{-\log P_{Y_i^n}(Y_i^n)}{\sqrt{n}} - \gamma_n < \frac{-\log P_{Y^n}(Y_i^n)}{\sqrt{n}} \right\} \geq 1 - e^{-\sqrt{n}\gamma_n},$$

holds for $i = 1, 2$. So we have

$$\frac{-\log P_{Y_i^n}(Y_i^n)}{\sqrt{n}} - \gamma_n < \frac{-\log P_{Y^n}(Y_i^n)}{\sqrt{n}}$$

with probability $1 - e^{-\sqrt{n}\gamma_n}$. So, we have for $i = 1, 2$

$$\Pr \left\{ \frac{-\log P_{Y^n}(Y_i^n)}{\sqrt{n}} \geq z_n \right\} \geq \Pr \left\{ \frac{-\log P_{Y_i^n}(Y_i^n)}{\sqrt{n}} - \gamma_n \geq z_n \right\} - e^{-\sqrt{n}\gamma_n},$$

which is the first inequality of the lemma. Secondly, we show the second inequality of the lemma.

Set

$$\begin{aligned} S_n(z_n) &= \left\{ \mathbf{y} \in \mathcal{Y}^n \mid P_{Y^n}(\mathbf{y}) \leq e^{-\sqrt{n}z_n} \right\} \\ S_n^{(i)}(z_n) &= \left\{ \mathbf{y} \in \mathcal{Y}^n \mid P_{Y_i^n}(\mathbf{y}) \leq \frac{e^{-\sqrt{n}z_n}}{w(i)} \right\} \end{aligned}$$

From the property of mixed sources, $\mathbf{y} \in S_n^{(i)}(z_n)$ ($i = 1, 2$) holds for $\mathbf{y} \in S_n(z_n)$. This means that

$$S_n(z_n) \subset S_n^{(i)}(z_n)$$

($i = 1, 2$). Moreover, since, by assumption, $\gamma_n \geq \frac{-\log w(i)}{\sqrt{n}}$ ($i = 1, 2$) hold for sufficiently large n , we have

$$\begin{aligned} &\Pr \left\{ \frac{-\log P_{Y^n}(Y_i^n)}{\sqrt{n}} \geq z_n \right\} \\ &= \Pr \left\{ P_{Y^n}(Y_i^n) \leq e^{-\sqrt{n}z_n} \right\} \\ &= P_{Y_i^n}(S_n(z_n)) \\ &\leq P_{Y_i^n}(S_n^{(i)}(z_n)) \\ &= \Pr \left\{ \frac{-\log P_{Y_i^n}(Y_i^n)}{\sqrt{n}} \geq z_n - \frac{-\log w(i)}{\sqrt{n}} \right\} \\ &\leq \Pr \left\{ \frac{-\log P_{Y_i^n}(Y_i^n)}{\sqrt{n}} \geq z_n - \gamma_n \right\}, \end{aligned}$$

for sufficiently large n , which is the second inequality. ■

APPENDIX C

PROOF OF THEOREM 5.1

Proof of Case I:

A simplest way to prove Theorem 5.1 is to first apply Theorem 3.1 to the present case of mixed sources and to proceed to the computation of necessary quantities. Here, however, more basically we start along with Lemma 5.1 and Lemma 5.2 in order to reveal the fundamental logic underlying the whole process of random number generation. Actually, the computations needed in the first way of proof are contained in those needed in the second way of proof; more exactly, both computations are substantially the same.

We show the proof of Case I by using Lemma 5.1 and Lemma 5.2. The proof consists of two parts.

1) Direct Part:

Set $M_n = e^{nH(Y_1) + T_1\sqrt{n} + \gamma}$, where $\gamma > 0$ is an arbitrarily small number. Then, trivially it holds that

$$\limsup_{n \rightarrow \infty} \frac{\log M_n - nH(Y_1)}{\sqrt{n}} \leq T_1.$$

Thus, it is enough to show that there exists a mapping ϕ_n such that

$$\limsup_{n \rightarrow \infty} d(\phi_n(U_{M_n}), Y^n) \leq \delta.$$

On the other hand, set $z_n = \frac{nH(Y_1) + T_1\sqrt{n}}{\sqrt{n}} - \gamma$, then $z_n + \gamma \leq \frac{\log M_n}{\sqrt{n}}$ holds. Thus, from Lemma 5.1 and (5.3) there exists a mapping ϕ_n such that

$$\begin{aligned} & \frac{1}{2} d(\phi_n(U_{M_n}), Y^n) \\ & \leq \Pr \left\{ \frac{nH(Y_1) + T_1\sqrt{n}}{\sqrt{n}} - \gamma < \frac{1}{\sqrt{n}} \log \frac{1}{P_{Y^n}(Y^n)} \right\} + e^{-\sqrt{n}\gamma}. \end{aligned}$$

Moreover, from Lemma 4.1, there exists a mapping ϕ_n such that

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \frac{1}{2} d(\phi_n(U_{M_n}), Y^n) \\ & \leq \limsup_{n \rightarrow \infty} \Pr \left\{ \frac{nH(Y_1) + T_1\sqrt{n}}{\sqrt{n}} - \gamma < \frac{1}{\sqrt{n}} \log \frac{1}{P_{Y^n}(Y^n)} \right\} \\ & = \limsup_{n \rightarrow \infty} \sum_{i=1}^2 \Pr \left\{ \frac{nH(Y_1) + T_1\sqrt{n}}{\sqrt{n}} - \gamma < \frac{1}{\sqrt{n}} \log \frac{1}{P_{Y^n}(Y_i^n)} \right\} w(i) \end{aligned}$$

$$\begin{aligned}
&\leq \sum_{i=1}^2 \limsup_{n \rightarrow \infty} \Pr \left\{ \frac{nH(Y_1) + T_1 \sqrt{n}}{\sqrt{n}} - \gamma - \gamma_n < \frac{1}{\sqrt{n}} \log \frac{1}{P_{Y_i^n}(Y_i^n)} \right\} w(i) \\
&\leq \sum_{i=1}^2 \limsup_{n \rightarrow \infty} \Pr \left\{ \frac{-\log P_{Y_i^n}(Y_i^n) - nH(Y_1)}{\sqrt{n}} > T_1 - 2\gamma \right\} w(i),
\end{aligned}$$

because $\gamma_n > 0$ is specified in Lemma 4.1, and so $\gamma_n < \gamma$ holds for sufficiently large n . Then, noting that $H(Y_1) = H(Y_2)$ holds, we have from the asymptotic normality (by virtue of the central limit theorem)

$$\begin{aligned}
&\limsup_{n \rightarrow \infty} \Pr \left\{ \frac{-\log P_{Y_i^n}(Y_i^n) - nH(Y_1)}{\sqrt{n}\sigma_i} > \frac{T_1 - 2\gamma}{\sigma_i} \right\} \\
&= \int_{\frac{T_1 - 2\gamma}{\sigma_i}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{z^2}{2} \right] dz \\
&= \int_{\frac{T_1}{\sigma_i}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{z^2}{2} \right] dz + \int_{\frac{T_1 - 2\gamma}{\sigma_i}}^{\frac{T_1}{\sigma_i}} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{z^2}{2} \right] dz.
\end{aligned}$$

Here, by the continuity of the normal distribution,

$$\int_{\frac{T_1 - 2\gamma}{\sigma_i}}^{\frac{T_1}{\sigma_i}} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{z^2}{2} \right] dz \rightarrow 0$$

as $\gamma \rightarrow 0$. Thus, noting that $\gamma > 0$ is an arbitrarily small, we have

$$\begin{aligned}
&\limsup_{n \rightarrow \infty} \frac{1}{2} d(\phi_n(U_{M_n}), Y^n) \\
&\leq \sum_{i=1}^2 w(i) \int_{\frac{T_1}{\sigma_i}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{z^2}{2} \right] dz = \frac{\delta}{2}.
\end{aligned}$$

Therefore, Direct Part has been proved.

2) Converse Part:

We consider a constant $T'_1 < T_1$. Assuming that T'_1 is $(H(Y_1), \delta)$ -achievable, we shall show a contradiction. Since we assume that T'_1 is $(H(Y_1), \delta)$ -achievable, there exists a mapping ϕ_n such that

$$\begin{aligned}
&\limsup_{n \rightarrow \infty} d(\phi_n(U_{M_n}), Y^n) \leq \delta, \\
&\limsup_{n \rightarrow \infty} \frac{\log M_n - nH(Y_1)}{\sqrt{n}} \leq T'_1,
\end{aligned}$$

which means that there exists a constant $\gamma > 0$ satisfying

$$\frac{\log M_n - nH(Y_1)}{\sqrt{n}} < T'_1 + \gamma < T_1,$$

for sufficiently large n .

On the other hand, set $z_n = \frac{nH(Y_1) + T'_1\sqrt{n}}{\sqrt{n}} + \gamma$. Then $z_n > \frac{\log M_n}{\sqrt{n}}$ holds. Thus, from Lemma 5.2 and (5.3), for any mapping ϕ_n it holds that

$$\begin{aligned} & \frac{1}{2}d(\phi_n(U_{M_n}), Y^n) \\ & \geq \Pr \left\{ \frac{nH(Y_1) + T'_1\sqrt{n}}{\sqrt{n}} + 2\gamma < \frac{1}{\sqrt{n}} \log \frac{1}{P_{Y^n}(Y^n)} \right\} - e^{-\sqrt{n}\gamma}. \end{aligned}$$

Thus, from Lemma 4.1, for any mapping ϕ_n we have

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \frac{1}{2}d(\phi_n(U_{M_n}), Y^n) \\ & \geq \liminf_{n \rightarrow \infty} \Pr \left\{ \frac{nH(Y_1) + T'_1\sqrt{n}}{\sqrt{n}} + 2\gamma < \frac{1}{\sqrt{n}} \log \frac{1}{P_{Y^n}(Y^n)} \right\} \\ & = \liminf_{n \rightarrow \infty} \sum_{i=1}^2 \Pr \left\{ \frac{nH(Y_1) + T'_1\sqrt{n}}{\sqrt{n}} + 2\gamma < \frac{1}{\sqrt{n}} \log \frac{1}{P_{Y^n}(Y_i^n)} \right\} w(i) \\ & \geq \sum_{i=1}^2 \liminf_{n \rightarrow \infty} \Pr \left\{ \frac{nH(Y_1) + T'_1\sqrt{n}}{\sqrt{n}} + 2\gamma + \gamma_n < \frac{1}{\sqrt{n}} \log \frac{1}{P_{Y_i^n}(Y_i^n)} \right\} w(i) \\ & = \sum_{i=1}^2 \liminf_{n \rightarrow \infty} \Pr \left\{ \frac{-\log P_{Y_i^n}(Y_i^n) - nH(Y_1)}{\sqrt{n}} > T'_1 + 2\gamma + \gamma_n \right\} w(i) \\ & \geq \sum_{i=1}^2 \liminf_{n \rightarrow \infty} \Pr \left\{ \frac{-\log P_{Y_i^n}(Y_i^n) - nH(Y_1)}{\sqrt{n}} > T'_1 + 3\gamma \right\} w(i), \end{aligned} \tag{C.1}$$

because $\gamma > \gamma_n$ holds for sufficiently large n (cf. Lemma 4.1). Then, by virtue of the asymptotic normality, for $i = 1, 2$ it holds that

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \Pr \left\{ \frac{-\log P_{Y_i^n}(Y_i^n) - nH(Y_1)}{\sqrt{n}\sigma_i} > \frac{T'_1 + 3\gamma}{\sigma_i} \right\} \\ & = \int_{\frac{T'_1 + 3\gamma}{\sigma_i}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{z^2}{2} \right] dz \\ & > \int_{\frac{T_1}{\sigma_i}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{z^2}{2} \right] dz, \end{aligned}$$

because we can let $T'_1 + 3\gamma < T_1$ hold by letting $\gamma \rightarrow 0$, and by substituting the above inequality into (C.1) for any mapping ϕ_n , we have

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \frac{1}{2}d(\phi_n(U_{M_n}), Y^n) \\ & > \sum_{i=1}^2 w(i) \int_{\frac{T_1}{\sigma_i}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{z^2}{2} \right] dz = \frac{\delta}{2}. \end{aligned}$$

This is a contradiction. Therefore, the converse part has been proved. ■

Proofs of Case II and Case III:

The proofs of Case II and Case III are similar to the proof of Case I. Here, we show only differences in the proofs.

1) *Direct Part:*

a) Similarly to *Case I*: we have the proof of *Case II* as follows.

$$\begin{aligned}
& \limsup_{n \rightarrow \infty} \frac{1}{2} d(\phi_n(U_{M_n}), Y^n) \\
& \leq \sum_{i=1}^2 \limsup_{n \rightarrow \infty} \Pr \left\{ \frac{-\log P_{Y_i^n}(Y_i^n) - nH(Y_1)}{\sqrt{n}} > T_2 - 2\gamma \right\} w(i) \\
& = \sum_{i=1}^2 \limsup_{n \rightarrow \infty} \Pr \left\{ \frac{-\log P_{Y_i^n}(Y_i^n) - nH(Y_1)}{\sqrt{n}\sigma_i} > \frac{T_2 - 2\gamma}{\sigma_i} \right\} w(i) \\
& = \limsup_{n \rightarrow \infty} \Pr \left\{ \frac{-\log P_{Y_1^n}(Y_1^n) - nH(Y_1)}{\sqrt{n}\sigma_1} > \frac{T_2 - 2\gamma}{\sigma_1} \right\} w(1) \\
& \quad + \limsup_{n \rightarrow \infty} \Pr \left\{ \frac{-\log P_{Y_2^n}(Y_2^n) - nH(Y_2)}{\sqrt{n}\sigma_2} > \frac{n(H(Y_1) - H(Y_2))}{\sqrt{n}\sigma_2} + \frac{T_2 - 2\gamma}{\sigma_2} \right\} w(2).
\end{aligned}$$

Here, note that $H(Y_1) > H(Y_2)$ holds. This means that for any constant $W_1 > 0$

$$\frac{n(H(Y_1) - H(Y_2))}{\sqrt{n}\sigma_2} + \frac{T_2 - 2\gamma}{\sigma_2} > W_1,$$

holds for sufficiently large n . Thus, taking account of $H(Y_1) > H(Y_2)$, we have

$$\begin{aligned}
& \limsup_{n \rightarrow \infty} \Pr \left\{ \frac{-\log P_{Y_2^n}(Y_2^n) - nH(Y_2)}{\sqrt{n}\sigma_2} > \frac{n(H(Y_1) - H(Y_2))}{\sqrt{n}\sigma_2} + \frac{T_2 - 2\gamma}{\sigma_2} \right\} \\
& \leq \limsup_{n \rightarrow \infty} \Pr \left\{ \frac{-\log P_{Y_2^n}(Y_2^n) - nH(Y_2)}{\sqrt{n}\sigma_2} > W_1 \right\} \\
& = \int_{W_1}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{y^2}{2} \right] dy.
\end{aligned}$$

Since $W_1 > 0$ can be arbitrarily large, we have

$$\lim_{n \rightarrow \infty} \Pr \left\{ \frac{-\log P_{Y_2^n}(Y_2^n) - nH(Y_2)}{\sqrt{n}\sigma_2} > \frac{n(H(Y_1) - H(Y_2))}{\sqrt{n}\sigma_2} + \frac{T_2 - 2\gamma}{\sigma_2} \right\} = 0.$$

Thus, from the asymptotic normality it follows that

$$\begin{aligned}
& \sum_{i=1}^2 \limsup_{n \rightarrow \infty} \Pr \left\{ \frac{-\log P_{Y_i^n}(Y_i^n) - nH(Y_1)}{\sqrt{n}} > T_2 - 2\gamma \right\} w(i) \\
& = \limsup_{n \rightarrow \infty} \Pr \left\{ \frac{-\log P_{Y_1^n}(Y_1^n) - nH(Y_1)}{\sqrt{n}\sigma_1} > \frac{T_2 - 2\gamma}{\sigma_1} \right\} w(1) \\
& = w(1) \int_{\frac{T_2 - 2\gamma}{\sigma_1}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{z^2}{2} \right] dz
\end{aligned}$$

$$\rightarrow w(1) \int_{\frac{T_2}{\sigma_1}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{z^2}{2} \right] dz = \frac{\delta}{2},$$

by letting $\gamma \rightarrow 0$. Thus, we have proved the direct part of Case II.

b) In *Case III*: we have

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \frac{1}{2} d(\phi_n(U_{M_n}), Y^n) \\ & \leq \sum_{i=1}^2 \limsup_{n \rightarrow \infty} \Pr \left\{ \frac{-\log P_{Y_i^n}(Y_i^n) - nH(Y_2)}{\sqrt{n}} > T_3 - 2\gamma \right\} w(i) \\ & \leq w(1) + \limsup_{n \rightarrow \infty} \Pr \left\{ \frac{-\log P_{Y_2^n}(Y_2^n) - nH(Y_2)}{\sqrt{n}\sigma_2} > \frac{T_3 - 2\gamma}{\sigma_2} \right\} w(2) \\ & \rightarrow w(1) + w(2) \int_{\frac{T_3}{\sigma_2}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{z^2}{2} \right] dz = \frac{\delta}{2} \end{aligned}$$

by letting $\gamma \rightarrow 0$. Consequently, similarly to the proof of Case I or Case II, we can prove the direct part of Case III.

2) *Converse Part*:

a) In *Case II*: we consider a constant $T'_2 < T_2$ and we assume that T'_2 is $(H(Y_1), \delta)$ -achievable. Notice that there exists a constant $\gamma > 0$ satisfying $T'_2 + \gamma < T_2$. Then, similarly to the proof of Case I, we have

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \frac{1}{2} d(\phi_n(U_{M_n}), Y^n) \\ & \geq \sum_{i=1}^2 \liminf_{n \rightarrow \infty} \Pr \left\{ \frac{-\log P_{Y_i^n}(Y_i^n) - nH(Y_1)}{\sqrt{n}} > T'_2 + 2\gamma + \gamma_n \right\} w(i) \\ & \geq \liminf_{n \rightarrow \infty} \Pr \left\{ \frac{-\log P_{Y_1^n}(Y_1^n) - nH(Y_1)}{\sqrt{n}} > T'_2 + 3\gamma \right\} w(1) \\ & = w(1) \int_{\frac{T'_2 + 3\gamma}{\sigma_1}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{z^2}{2} \right] dz \\ & > w(1) \int_{\frac{T_2}{\sigma_1}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{z^2}{2} \right] dz = \frac{\delta}{2}, \end{aligned}$$

because we can let $T'_2 + 3\gamma < T_2$ by letting $\gamma \rightarrow 0$. Hence, we have shown the converse part by using the similar argument to Case I.

b) Similarly, in *Case III*: we consider a constant $T'_3 < T_3$ and we assume that T'_3 is $(H(Y_2), \delta)$ -

achievable. Notice that there exists a constant $\gamma > 0$ satisfying $T'_3 + \gamma < T_3$. Then, we have

$$\begin{aligned}
& \liminf_{n \rightarrow \infty} \frac{1}{2} d(\phi_n(U_{M_n}), Y^n) \\
& \geq \sum_{i=1}^2 \liminf_{n \rightarrow \infty} \Pr \left\{ \frac{-\log P_{Y_i^n}(Y_i^n) - nH(Y_2)}{\sqrt{n}} > T'_3 + 2\gamma + \gamma_n \right\} w(i) \\
& \geq \liminf_{n \rightarrow \infty} \Pr \left\{ \frac{-\log P_{Y_1^n}(Y_1^n) - nH(Y_1)}{\sqrt{n}\sigma_1} > \frac{n(H(Y_2) - H(Y_1))}{\sqrt{n}\sigma_1} + \frac{T'_3 + 3\gamma}{\sigma_1} \right\} w(1) \\
& \quad + \liminf_{n \rightarrow \infty} \Pr \left\{ \frac{-\log P_{Y_2^n}(Y_2^n) - nH(Y_2)}{\sqrt{n}\sigma_2} > \frac{T'_3 + 3\gamma}{\sigma_2} \right\} w(2). \tag{C.2}
\end{aligned}$$

Then, the first term of the right-hand side of the above inequality can be determined as follows.

Notice that $H(Y_1) > H(Y_2)$ holds. This means that for any constant $W_1 > 0$

$$\frac{n(H(Y_2) - H(Y_1))}{\sqrt{n}\sigma_1} + \frac{T'_3 + 3\gamma}{\sigma_1} < -W_1,$$

holds for sufficiently large n . Thus,

$$\begin{aligned}
& \liminf_{n \rightarrow \infty} \Pr \left\{ \frac{-\log P_{Y_1^n}(Y_1^n) - nH(Y_1)}{\sqrt{n}\sigma_1} > \frac{n(H(Y_2) - H(Y_1))}{\sqrt{n}\sigma_1} + \frac{T'_3 + 3\gamma}{\sigma_1} \right\} \\
& \geq \liminf_{n \rightarrow \infty} \Pr \left\{ \frac{-\log P_{Y_1^n}(Y_1^n) - nH(Y_1)}{\sqrt{n}\sigma_1} > -W_1 \right\} \\
& = \int_{-W_1}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{y^2}{2} \right] dy,
\end{aligned}$$

holds. Since $W_1 > 0$ can be arbitrarily large, we have

$$\lim_{n \rightarrow \infty} \Pr \left\{ \frac{-\log P_{Y_1^n}(Y_1^n) - nH(Y_1)}{\sqrt{n}\sigma_1} > \frac{n(H(Y_2) - H(Y_1))}{\sqrt{n}\sigma_1} + \frac{T'_3 + 3\gamma}{\sigma_1} \right\} = 1.$$

Substituting the above equality into (C.2), by virtue of the asymptotic normality, it holds that

$$\begin{aligned}
& \liminf_{n \rightarrow \infty} \frac{1}{2} d(\phi_n(U_{M_n}), Y^n) \\
& \geq \sum_{i=1}^2 \liminf_{n \rightarrow \infty} \Pr \left\{ \frac{-\log P_{Y_i^n}(Y_i^n) - nH(Y_2)}{\sqrt{n}} > T'_3 + 3\gamma \right\} w(i) \\
& = w(1) + \liminf_{n \rightarrow \infty} \Pr \left\{ \frac{-\log P_{Y_2^n}(Y_2^n) - nH(Y_2)}{\sqrt{n}\sigma_2} > \frac{T'_3 + 3\gamma}{\sigma_2} \right\} w(2) \\
& = w(1) + w(2) \int_{\frac{T'_3 + 3\gamma}{\sigma_1}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{z^2}{2} \right] dz \\
& > w(1) + w(2) \int_{\frac{T_3}{\sigma_1}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{z^2}{2} \right] dz = \frac{\delta}{2},
\end{aligned}$$

because we can let $T'_3 + 3\gamma < T_3$ by letting $\gamma \rightarrow 0$. Hence, we have shown the proofs similarly to

Case I. ■

APPENDIX D

PROOF OF THEOREM 7.1

1) *Direct Part:*

We prove that $L_1 = T_7 + \gamma$ is an $(H(Y_1), \varepsilon)$ -achievable, where $\gamma > 0$ is an arbitrary small constant. Set $M_n = e^{nH(Y_1) + \sqrt{n}L_1}$. Then, obviously we have

$$\limsup_{n \rightarrow \infty} \frac{\log M_n - nH(Y_1)}{\sqrt{n}} \leq L_1.$$

Thus, it is enough to show that there exists an (n, M_n, ε_n) code satisfying

$$\limsup_{n \rightarrow \infty} \varepsilon_n \leq \varepsilon.$$

From Lemma 7.1, there exists an (n, M_n, ε_n) code that satisfies

$$\begin{aligned} \varepsilon_n &\leq \Pr \left\{ \frac{1}{\sqrt{n}} \log \frac{1}{P_{Y^n}(Y^n)} \geq \frac{1}{\sqrt{n}} \log M_n \right\} \\ &= \Pr \left\{ \frac{1}{\sqrt{n}} \log \frac{1}{P_{Y^n}(Y^n)} \geq \frac{nH(Y_1) + \sqrt{n}L_1}{\sqrt{n}} \right\} \\ &= \Pr \left\{ \frac{-\log P_{Y^n}(Y^n) - nH(Y_1)}{\sqrt{n}} \geq L_1 \right\} \\ &= \sum_{i=1}^2 \Pr \left\{ \frac{-\log P_{Y^n}(Y_i^n) - nH(Y_1)}{\sqrt{n}} \geq L_1 \right\} w(i). \end{aligned} \quad (\text{D.1})$$

The last equality is derived from the definition of the mixed source. Then, from Lemma 4.1, we have

$$\begin{aligned} &\limsup_{n \rightarrow \infty} \sum_{i=1,2} \Pr \left\{ \frac{-\log P_{Y^n}(Y_i^n) - nH(Y_1)}{\sqrt{n}} \geq L_1 \right\} w(i) \\ &\leq \sum_{i=1}^2 \limsup_{n \rightarrow \infty} \Pr \left\{ \frac{-\log P_{Y_i^n}(Y_i^n) - nH(Y_1)}{\sqrt{n}} \geq L_1 - \gamma_n \right\} w(i) \\ &\leq \sum_{i=1}^2 \limsup_{n \rightarrow \infty} \Pr \left\{ \frac{-\log P_{Y_i^n}(Y_i^n) - nH(Y_1)}{\sqrt{n}\sigma_i} \geq \frac{L_1 - \gamma}{\sigma_i} \right\} w(i) \\ &= \sum_{i=1}^2 \limsup_{n \rightarrow \infty} \Pr \left\{ \frac{-\log P_{Y_i^n}(Y_i^n) - nH(Y_1)}{\sqrt{n}\sigma_i} \geq \frac{T_7}{\sigma_i} \right\} w(i), \end{aligned}$$

because γ_n is specified in Lemma 4.1, and so $\gamma > \gamma_n$ holds for sufficiently large n .

Then, noting that $H(Y_1) = H(Y_2)$ holds, from the asymptotic normality, we have

$$\limsup_{n \rightarrow \infty} \Pr \left\{ \frac{-\log P_{Y_i^n}(Y_i^n) - nH(Y_1)}{\sqrt{n}\sigma_i} \geq \frac{T_7}{\sigma_i} \right\} = \int_{\frac{T_7}{\sigma_i}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{z^2}{2} \right] dz,$$

for $i = 1, 2$. Thus, we have

$$\limsup_{n \rightarrow \infty} \sum_{i=1}^2 \Pr \left\{ \frac{-\log P_{Y^n}(Y_i^n) - nH(Y_1)}{\sqrt{n}} \geq L_1 \right\} w(i) \leq \sum_{i=1}^2 w(i) \int_{\frac{T_7}{\sigma_i}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{z^2}{2} \right] dz.$$

Substituting the above inequality into (D.1), we have

$$\limsup_{n \rightarrow \infty} \epsilon_n \leq \sum_{i=1}^2 w(i) \int_{\frac{T_7}{\sigma_i}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{z^2}{2} \right] dz = \varepsilon,$$

where the last equality is derived from the definition of T_7 given by (7.2). Therefore, Direct Part has been proved.

2) Converse Part:

Let us assume that L_2 satisfying $L_2 < T_7$ is $(H(Y_1), \varepsilon)$ -achievable. Then we shall show a contradiction. Notice that there exists a constant $\gamma > 0$ such that $L_2 + 3\gamma < T_7$ holds.

Then from the assumption there must exist an (n, M_n, ϵ_n) code such that

$$\begin{aligned} \limsup_{n \rightarrow \infty} \epsilon_n &\leq \varepsilon, \\ \limsup_{n \rightarrow \infty} \frac{\log M_n - nH(Y_1)}{\sqrt{n}} &\leq L_2, \end{aligned}$$

This means that for sufficiently large n , there must exist an (n, M_n, ϵ_n) code such that

$$\frac{\log M_n}{\sqrt{n}} < \frac{nH(Y_1)}{\sqrt{n}} + L_2 + \gamma. \quad (\text{D.2})$$

for any $\gamma > 0$. On the other hand, from Lemma 7.2, any (n, M_n, ϵ_n) code satisfies

$$\epsilon_n \geq \Pr \left\{ \frac{1}{\sqrt{n}} \log \frac{1}{P_{Y^n}(Y^n)} \geq \frac{1}{\sqrt{n}} \log M_n + \gamma \right\} - e^{-\sqrt{n}\gamma}.$$

Thus, from (D.2) it follows that

$$\begin{aligned} \epsilon_n &\geq \Pr \left\{ \frac{1}{\sqrt{n}} \log \frac{1}{P_{Y^n}(Y^n)} \geq \frac{1}{\sqrt{n}} \log M_n + \gamma \right\} - e^{-\sqrt{n}\gamma} \\ &\geq \Pr \left\{ \frac{1}{\sqrt{n}} \log \frac{1}{P_{Y^n}(Y^n)} \geq \frac{nH(Y_1)}{\sqrt{n}} + L_2 + 2\gamma \right\} - e^{-\sqrt{n}\gamma} \\ &= \sum_{i=1}^2 \Pr \left\{ \frac{-\log P_{Y^n}(Y_i^n) - nH(Y_1)}{\sqrt{n}} \geq L_2 + 2\gamma \right\} w(i) - e^{-\sqrt{n}\gamma}, \end{aligned}$$

for sufficiently large n , where the last equality is derived from the definition of the mixed source.

From Lemma 4.1, we have

$$\begin{aligned}
\liminf_{n \rightarrow \infty} \varepsilon_n &\geq \liminf_{n \rightarrow \infty} \sum_{i=1}^2 \Pr \left\{ \frac{-\log P_{Y^n}(Y_i^n) - nH(Y_1)}{\sqrt{n}} \geq L_2 + 2\gamma \right\} w(i) \\
&\geq \sum_{i=1}^2 \liminf_{n \rightarrow \infty} \Pr \left\{ \frac{-\log P_{Y^n}(Y_i^n) - nH(Y_1)}{\sqrt{n}} \geq L_2 + 2\gamma + \gamma_n \right\} w(i) \\
&\geq \sum_{i=1}^2 \liminf_{n \rightarrow \infty} \Pr \left\{ \frac{-\log P_{Y^n}(Y_i^n) - nH(Y_1)}{\sqrt{n}} \geq L_2 + 3\gamma \right\} w(i), \tag{D.3}
\end{aligned}$$

because γ_n is specified in Lemma 4.1 and so $\gamma_n < \gamma$ holds for sufficiently large n . Then, by virtue of the asymptotic normality we have

$$\begin{aligned}
&\sum_{i=1}^2 \liminf_{n \rightarrow \infty} \Pr \left\{ \frac{-\log P_{Y^n}(Y_i^n) - nH(Y_1)}{\sqrt{n}} \geq L_2 + 3\gamma \right\} w(i) \\
&= \sum_{i=1}^2 \liminf_{n \rightarrow \infty} \Pr \left\{ \frac{-\log P_{Y^n}(Y_i^n) - nH(Y_1)}{\sqrt{n}\sigma_i} \geq \frac{L_2 + 3\gamma}{\sigma_i} \right\} w(i) \\
&= \sum_{i=1}^2 w(i) \int_{\frac{L_2 + 3\gamma}{\sigma_i}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{z^2}{2} \right] dz \\
&> \sum_{i=1}^2 w(i) \int_{\frac{T_7}{\sigma_i}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{z^2}{2} \right] dz = \varepsilon,
\end{aligned}$$

because we can let $L_2 + 3\gamma < T_7$ holds by letting $\gamma \rightarrow 0$ and the last equality is from (7.2). Thus, substituting the above into (D.3), it must hold that

$$\liminf_{n \rightarrow \infty} \varepsilon_n > \varepsilon.$$

This is a contradiction and we conclude that L_2 satisfying $L_2 < T_7$ is not an $(H(Y_1), \varepsilon)$ -achievable.

■

Proofs of Case II and Case III:

The proofs of Case II and Case III can be shown by the same arguments as the proof of Case I of Theorem 7.1. ■

ACKNOWLEDGMENT

The authors are very grateful to the anonymous reviewers for their helpful comments, which have occasioned to improve the main results so as to be much more general and transparent. The first author is very grateful to Prof. Toshiyasu Matsushima of Waseda University for his valuable discussions and comments.

REFERENCES

- [1] T. S. Han and S. Verdú, "Approximation theory of output statistics," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 752–772, 1993.
- [2] S. Vembu and S. Verdú, "Generating random bits from an arbitrary source: Fundamental limits," *IEEE Trans. Inf. Theory*, vol. 41, no. 5, pp. 1322–1332, 1995.
- [3] T. S. Han, "Folklore in source coding: Information-spectrum approach," *IEEE Trans. Inf. Theory*, vol. 51, no. 2, pp. 747–753, 2005.
- [4] K. Visweswariah, S. R. Kulkarni, and S. Verdú, "Source codes as random number generators," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 462–471, 1998.
- [5] T. S. Han, *Information-Spectrum Methods in Information Theory*. Springer, New York, 2003.
- [6] Y. Steinberg and S. Verdú, "Simulation of random processes and rate-distortion theory," *IEEE Trans. Inf. Theory*, vol. 42, no. 1, pp. 63–86, 1996.
- [7] I. Kontoyiannis, "Second-order noiseless source coding theorems," *IEEE Trans. Inf. Theory*, vol. 43, no. 4, pp. 1339–1341, 1997.
- [8] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Academic, 1981.
- [9] M. Hayashi, "Information spectrum approach to second-order coding rate in channel coding," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 4947–4966, 2009.
- [10] Y. Polyanskiy, H. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, 2010.
- [11] M. Hayashi, "Second-order asymptotics in fixed-length source coding and intrinsic randomness," *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4619–4637, 2008.
- [12] R. Nomura and T. Matsushima, "Achievable condition in resolvability problem for mixed sources," in *Proc. The 33rd Symposium on Information Theory and its Applications*, 2010, pp. 138–143.
- [13] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Dispersion of the Gilbert-Elliott Channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 1829–1848, 2011.
- [14] K. L. Chung, *Markov Chains with Stationary Transition Probabilities, 2nd edition*. Springer, New York, 1966.
- [15] P. Billingsley, *Probability and Measure*. Wiley, 1995.
- [16] P. C. Shields, *The Ergodic Theory of Discrete Sample Paths*. American Mathematical Society, 1996.
- [17] A. J. Goldsmith and P. P. Varaiya, "Capacity, mutual information, and coding for finite-state Markov channels," *IEEE Trans. Inf. Theory*, vol. 42, no. 3, pp. 868–886, 1996.
- [18] B. D. Fritchman, "A binary channel characterization using partitioned Markov chain," *IEEE Trans. Inf. Theory*, vol. IT-13, no. 2, pp. 221–227, 1967.
- [19] R. Nomura and T. S. Han, "Capacity theorems for the Fritchman channel," in preparation.